JAN
Leading Global Nursing Research WILEY

Short version of the Multidimensional Scale of Dating Violence (MSDV 2.0) in Spanish-language: Instrument development and psychometric evaluation

María de los Ángeles García-Carpintero-Muñoz 📵 | **Lorena Tarriño-Concejero** 📵 |
Eugenia Gil-García 📵 | **Ana María Pórcel-Gálvez** 📵 | **Sergio Barrientos-Trigo** 📵

Research Group PAIDI-CTS 1050 Complex Care, Chronicity and Health Outcomes, Faculty of Nursing, Physiotherapy and Podiatry, University of Seville, Seville, Spain

Correspondence
Lorena Tarriño-Concejero, Faculty of Nursing, Physiotherapy and Podiatry, University of Seville, C/ Avenzoar, 6. 41009 Seville, Spain.
Email: ltarrino@us.es

Abstract

Aims: To develop and psychometrically test the short version of the Multidimensional Scale of Dating Violence (MSDV 2.0) in Spanish-language to detect violence perpetrated and suffered in dating relationships.

Design: A psychometric instrument development and validation study.

Methods: A two-phase approach was used: Phase (1) the items of the original instrument were revised and new items related to online violence and sexual violence were incorporated. Content validation by a Delphi panel with 25 psychometric and dating violence experts were performed. Next, a face validity was performed in 32 students followed by a pilot study in another 74 participants. Phase (2) Psychometric validation, the instrument was tested in a sample of 1091 university students, analysing the psychometric properties based on construct validity and internal consistency. The study was conducted from September to November 2020 in the context of the Andalusian Public University System.

Results: In phase (1) 42 items for each subscale (perpetration, victimization) were accepted by the Delphi panel, and acceptable values were obtained for the criteria of clarity, coherence, and relevance. In phase (2) the MSDV 2.0 showed acceptable psychometric properties. Confirmatory factor analysis showed a five-dimensional structure with 18 items for each subscale with excellent fit rates. Reliability analysis indicated adequate internal consistency ($\alpha = .879-.802$) and correlations with the Depression, Anxiety, and Stress Scale ($\rho = .418-.225$) and the self-perceived health item ($\rho = .380-.179$), providing evidence of its convergent validity. Cut-off points were also calculated for each dimension, with their corresponding sensitivity and specificity, indicating to be a good instrument for detecting possible cases of dating violence.

Conclusion: The MSDV 2.0 is the only short instrument published to date that measures the dating violence suffered and perpetrated taking into account all its

dimensions. Its use would serve as support in prevention programs and design of public policies.

Impact: The short version of the MSDV 2.0 could be a comprehensive enough instrument to enable a detection and evaluation of dating violence in the educational setting.

KEYWORDS
assessments, nursing, adolescence and youth, dating violence, instrument development, intimate partner violence, nursing, psychometric properties, public health surveillance, validation studies

1 | INTRODUCTION

Dating violence (DV) is classified as a type of intimate partner violence (IPV) and is defined as the perpetration of emotional, physical or sexual abuse in a dating relationship, regardless of duration, and is an increasingly prevalent public health problem among adolescents and young adults (Emelianchik et al., 2018). However, it has not received as much attention in research as adult IPV (Straus & Gozjolko, 2014; United Nations Entity for Gender Equality and the Empowerment of Women [UN Women], 2021).

It is important to note that the victims of these behaviours are usually women. Globally, one in three women aged 15–49 years report having suffered some type of physical and/or sexual violence by their partner [World Health Organization (WHO, 2021)]. These data coincide with the latest study conducted in Europe with all member countries (European Union Agency for Fundamental Rights, 2014) and with the recent study conducted in Spain, the Macrosurvey on Violence against Women, which finds that psychological violence is suffered mainly by young women (38.3%) in the 16–24 years age bracket (Government Delegation against Gender Violence, Spain, 2019).

1.1 | Background

1.1.1 | The complexity of dating violence and its impact on health

In adolescence and youth, in the cultural construct of 'romantic love', objectionable behaviours such as jealousy or control over the partner, are regarded as positive elements and even evidence of love, which are idealized in couple relationships at a young age (Bosch et al., 2019). In addition, in many of these couples there is a normalization of violence as a strategy to resolve conflicts, becoming a habitual element in the relationship, which go unnoticed and are not detected as violence, since only physical aggressions are identified as violence (Perles et al., 2019; Rodríguez-Santero et al., 2017).

Violence in a dating relationship can include different acts, ranging from emotional abuse to physical aggression and/or sexual violence.

Regarding psycho-emotional abuse attitudes, we find actions of denigration and devaluation that set out to cause feelings of insecurity, such as threatening to break up with the other person or telling him/her that another person would be a better partner, ridiculing, insulting the other partner in front of others or telling him/her that he/she is a failure (García-Carpintero et al., 2018). In physical aggression behaviours, there is a wide range that escalates from threats or environmental violence to physical aggression (Rubio et al., 2015). The sexual violence dimension would include forced sexual relations by the partner through coercion with emotional control tactics, to forced sexual activity with or without penetration (Bagwell, 2021). In addition, having had multiple partners and/or consuming alcohol and other substances before the sexual act is associated with a greater propensity to be victims of sexual violence (Demissie et al., 2018; Khanhkham et al., 2020).

Moreover, information and communication technologies in social relationships plays a very important role in young people, as they spend most of their time using them (Gabelas & Marta, 2020). In some dating relationships, attitudes of control and surveillance can occur, including actions such as spying on photos and comments uploaded to networks, monitoring friendships or demanding that the partner delete content or profiles, monitoring the connection time or geolocation, thus creating a situation of effective control (Estébanez, 2018; Rodríguez et al., 2018).Cyberbullying is also found, which includes acts such as sending threats, disseminating private information and insulting a person on social media (Cava et al., 2020).

All these types of violent relationships generate negative consequences on health and lifestyle in adolescents and young people. Among others, high rates of low self-esteem, anxiety (Pérez-Marco et al., 2020), increase in eating disorders (Cha et al., 2016) and school failure (Banyard & Cross, 2008) are standing out consequences. It increases antisocial behaviours and risky behaviours, such as use of illicit substance, having multiple sex partners and engaging in unsafe sex, which can lead to sexually transmitted diseases and unwanted pregnancies (Kisa & Zeyneloğlu, 2019). All of them, in the long term, in their adulthood, could increase various mental disorders (Pérez-Marco et al., 2020), psychopathological disorders (Davila et al., 2016) and physical injury (Kisa & Zeyneloğlu, 2019). These consequences are more severe for girls

than for boys, implying greater physical and emotional sequelae (WHO, 2021). That is why health professionals, and especially nursing, have an outstanding role in the prevention and detection of DV (Jack et al., 2021).

1.1.2 | Validating measures of dating violence

Due to the high prevalence and incidence of DV, it is crucial that there be instruments that detect and measure these behaviours. Thus, over the last decade different instruments have been published that measure this phenomenon, encompassing several dimensions, such as the Dating Violence Questionnaire (DVQ; López et al., 2016), Violence in Adolescents' Dating Relationships Inventory (VADRI; Aizpitarte et al., 2017), Measure of Adolescent Relationship Harassment and Abuse (MARSHA; Rothman et al., 2021), Teen Dating Violence: Victimization and Perpetration scale. TDV-VP Scale (Soriano et al., 2021), Dating Violence Questionnaire-R (DVQ-R; Rodríguez-Díaz et al., 2017), Conflict in adolescent dating relationships inventory short form (CADRI-S; Fernández et al., 2012) and the Conflict Tactics Scales (CTS) in its different versions, modified (CTS-M; Ronzón et al., 2019) or revised (CTS-2; Anderson & Leigh, 2011). Most of these instruments have been validated in young university students, since at those ages, young people's first relationships usually begin or are consolidated (Llano-Suárez et al., 2021). Prevention and detection is also crucial in this context, since universities are institutions that promote health, ensure the right to the freedom, physical and emotional integrity of all its collective, and are guarantors of progress in our society (Spanish Network of Healthy Universities, 2021).

Designing valid and reliable instruments in the target population are a vital stage to be able to carry out the first step, which is to accurately detect the different types of violence and develop protocols or action plans to improve health in the adolescents and young people who are affected. In addition, the analysis of the data produced by these instruments can be used to develop prevention programs and public policies that promote the development of egalitarian relationships in young couples.

Our research team has validated the Multidimensional Scale of Dating Violence (MSDV; García-Carpintero et al., 2018). This instrument is composed of two subscales that explore the violence perpetrated and violence suffered. Each subscale contains 32 items with six answer options (0: never; 1: hardly ever; 2: more than three times; 3: more than four times; 4: more than 10 times; 5: always, usually). They measure six dimensions (physical and sexual abuse, harassment, surveillance, cyber bullying, domination and denigration) in the violence suffered and five dimensions (physical and sexual abuse, harassment, surveillance, domination and denigration) in the violence perpetrated. Internal consistency was adequate, with Cronbach's alpha between .93 and .91 for violence perpetrated and violence suffered, respectively, and with acceptable adjustment indices in the confirmatory factor analysis (CFA). Although it has proven to be valid and reliable, but we consider it pertinent to review

and update its items since, in recent years, online violence and sexual violence are marked by new behaviours.

On the other hand, the extensive literature on instruments that measure DV shows that there is still no brief instrument that assesses violence in young people in a comprehensive manner with an estimate of severity hence the relevance of our research. It is important to have a brief multidimensional instrument, which allows for easier use, with a cut-off score that can be used to establish a minimum threshold, which will follow a diagnostic approach and could serve as a community screening tool.

2 | THE STUDY

2.1 | Aims

The aim of the present study has been to develop the short version of the Multidimensional Scale of Dating Violence (MSDV 2.0) in Spanish-language and evaluate its psychometric properties.

2.2 | Design

A psychometric instrument development and validation methodological study were carried out in two phases: (1) design and content validation, and (2) psychometric validation. In the first phase, design and content validation of MSDV 2.0, the items of the original scale were revised and new items related to online violence and sexual violence were incorporated. Content validation, face validity and piloting were performed. In the second phase, psychometric validation, the scale was tested in a sample of university students, analysing the psychometric properties (construct validity and internal consistency) of the scale (Figure 1). The Consensus-based Standards for the Selection of Health Measurement Instruments (COSMIN) Study Design checklist was followed for this study (Mokkink et al., 2019). The study was conducted from September to November 2020.

2.3 | Instrument

The preliminary design of the instrument was developed in the first phase. The research team reviewed, formulated and incorporated new items based on the original scale (MSDV), and a review of the literature and instruments for measuring DV, focusing on the new behaviour related to sexual violence and online violence which has emerged in recent years. The whole process was carried out following the guidelines established by Morales et al. (2003). Of the 32 items per subscale (victimization and perpetration) of the MSDV, 13 items were reformulated, 10 new items were added and 19 items of the original scale were retained. A set of 42 items for each subscale was designed to cover five proposed dimensions (cyber bullying, control and surveillance, psycho-emotional, physical and sexual) basing our design on the MSDV and improving its structure by

FIGURE 1 Phases of the investigation
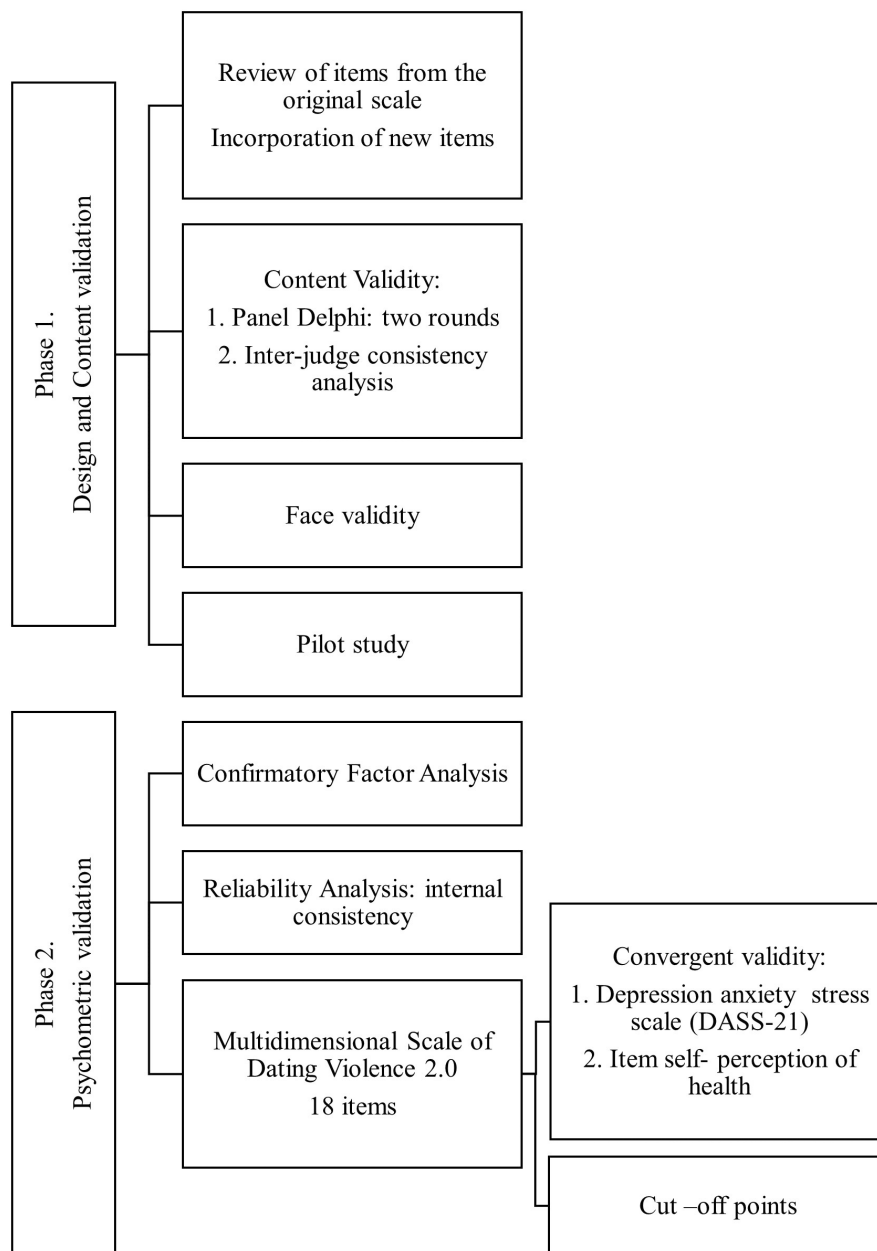
referring to the existing literature. Each item was accompanied by the phrase 'He/She has done it to me' for the victimization subscale and 'I have done it' for the perpetration subscale. They were scored on a Likert scale, with five response options (1: never; 2: sometimes (1 or 2 times); 3: occasionally (3–4 times); 4: repeatedly (5–10 times); 5: habitually (more than 10 times). The authors of the original scale and three researchers with experience in psychometrics and DV participated in this process.

Next, all the contents designed in the previous phase were subjected to a Delphi panel with the participation of 25 experts in DV and psychometrics. In the first round, 38 items were accepted for reaching mean scores between 3.12 and 3.96 in the criteria of clarity, coherence and relevance proposed by Abad et al. (2011) (accepting items with mean scores ≥3, in a range of 1–4, where the higher the score, the greater clarity, coherence and relevance) and an inter-rater agreement of content relevance, and a V Aiken

between .81 and .97 proposed by Penfield and Giacobbi (2004) (accepting items with values >0.7). Four items did not reach the minimum scores in one of the criteria: clarity. The responses to the open-ended questions and the suggestions provided were analysed. Modifications were made and proposed for a second round. In the second round, these four items were accepted because they achieved average values for the proposed criteria. The adequacy of the items for the proposed dimensions was also analysed, obtaining means between 3.44–3.72 and a V Aiken of 0.81–0.91. Finally, 42 items were obtained for five dimensions, by subscales. The whole process with the proposed items analysed in the first round and those reformulated in the second, as well as the scores, can be seen in Table A1.

In the face validity, 32 university students with an average age of 21.3 years (21 females and 11 males) participated. The objective was to determine the clarity, accuracy and comprehension of each of

the items (Cruz & Muñoz, 2015). No errors were identified, all items were understood, clear and accurate with mean scores of between 3.25 and 3.63, in a range of 1–4. The higher the score, the greater the clarity, accuracy and comprehension, and only scores ≥3 were accepted.

Finally, the Pilot test was carried out, which helps to reduce possible biases and errors in obtaining subsequent data and improves the methodology previously proposed (Van Van Teijlingen & Hundley, 2002). Seventy-four young university students participated in the study (67 women and seven men) of Spanish nationality, with a mean age of 19.59 years (*SD* 1.59). They were students on a Nursing Undergraduate Degree course at the University of Seville (51.4%) and the University of Jaen (48.6%). The mean time of the participants in a couple relationship was 28.23 months (*SD* 22.78) and 62% had a partner at the time of the study, while only one of the participants was co-habiting with their partner. No errors were detected related to the comprehension, acceptability and completion time of the questionnaire.

The 42 items by subscales belonging to five dimensions (Cyber bullying; Control and surveillance; Psycho-emotional; Physical and Sexual) produced by this phase were subjected to psychometric validation in the next phase.

2.4 | Sample and fieldwork

The second phase was conducted in the context of the Andalusian Public University System, composed of 10 state-funded universities with a total of 46,346 students enrolled in the 2019/2020 academic year (Integrated University Information System, 2021). A total of 1382 young university students were recruited by convenience from an awareness program funded by the Andalusian Youth Institute entitled 'Promoting healthy relationships in Andalusian university youth; prevention of gender violence', in which eight state-funded universities participated. Of the participants, 1091 young people met the inclusion criteria (having enrolled in the 2020/2021 academic year in an Andalusian university, being between 18 and 24 years old and having had or currently having a couple relationship).

2.5 | Data collection

Data were collected by means of a self-administered online questionnaire, via Google Forms. Sociodemographic variables were included: sex, age, nationality, university campus, university degree, average time in a relationship, whether they had a partner at the time of the study and whether they were living with a partner. The study was conducted from September to November 2020 in the context of the Andalusian Public University System.

For the DV analysis, the 42 items per subscale (victimization and perpetration) of the MSDV 2.0 resulting from the previous phase were incorporated.

For convergent validity, the Abbreviated Scale of Depression, Anxiety and Stress (DASS-21) scale and the self-perception of health item were used. The DASS-21 scale is a self-administered instrument with 21 items. It measures levels of anxiety, depression and stress, where higher scores correspond to worse levels in the respective dimensions. It has demonstrated good psychometric properties with an internal consistency of 0.9 for the total scale, 0.8 for the depression subscale, 0.73 for the anxiety subscale and 0.81 for the stress subscale. It has been validated in Spain in the university context by Fonseca et al. (2010). On the other hand, the item of self-perception of health describes health in general, where higher scores indicate a worse self-perception of health. It has been used in major international health surveys (Idler & Benyamini, 1997; Moreno et al., 2012) demonstrating its validity.

Dichotomous variables (yes/no) were also recorded for the validity of the criteria of the different dimensions for each subscale: Has your partner harassed you social media?/Have you harassed your partner on social media?; Has your partner controlled and/or monitored you?/Have you controlled and/or monitored your partner?; Have you felt denigrated and/or dominated by your partner?/Have you denigrated and/or dominated your partner?; Have you suffered any physical aggression from your partner?/Have you ever physically assaulted your partner?; Has your partner compelled or forced you to perform some sexual act that you did not want to do?/Have you compelled or forced your partner to perform some sexual act that he/she did not want to do?

2.6 | Ethical considerations

This research was approved by the Research Ethics Committee of the Virgen Macarena and Virgen del Rocío University Hospitals, study code VNRS_18. The young people participated voluntarily, received information about the study and signed an informed consent form. All data were processed safeguarding the confidentiality of the participants. The entire study complied with Spanish data protection legislation (Organic Law 3/, 2018).

2.7 | Data analysis

SPSS 26.0 (IBM Corp, 2019) and AMOS 23.0 (Arbuckle, 2014) programs were used for data analysis. Univariate analysis was performed by calculating absolute and relative frequencies, means and standard deviations. Ceiling and floor effects were calculated for all items for each subscale. Prior to the bivariate analysis, the Kolmogorov–Smirnov test was performed to test the normality of quantitative variables, which turned out not to follow normality. After that, the Mann–Whitney *U* test was used to check the association between dichotomous qualitative variables and continuous variables. The effect size for Mann–Whitney *U* test was calculated using the probability coefficient of superiority [*PSest*] estimating the following categories: No effect (PSest ≤0.0); small

(PSest ≤0.56); medium (PSest: 0.57–0.70) and large (PSest ≥0.71; Grissom, 1994).

2.7.1 | Validity analysis

Confirmatory factor analysis was performed to determine whether the data fit the proposed model after content validation on the two subscales. The method chosen was that of Robust maximum-likelihood estimation, to mitigate the possible biases that could occur in the estimates due to the observed floor effects (Brown, 2015). The following fit indices (FI) were used: (a) chi-square significance (χ^2), considered adequate if it was not significant for the model (Wang et al., 2017); (b) comparative fit index (CFI) and Tucker–Lewis Index (TLI), for which values >0.90 indicate an adequate fit, while >0.95 indicate an excellent fit (Bentler, 1990; Tucker & Lewis, 1973) and (c) root mean square error of approximation (RMSEA), for which values below 0.08 indicate an acceptable model fit and 0.06 an excellent fit, for the value and its 90% confidence interval (Hu & Bentler, 2009).

We considered fitting the instrument by studying the discrepancies between the model and the data through the modification indexes (MI; Sörbom, 1989). MI indicates the amount of decrease in the chi-square statistic of the model with one degree of freedom. A high value of MI indicates that the corresponding fixed parameter should be released, by removing it, to improve the model fit (Wang et al., 2017). Therefore, specific items with higher MI were removed, which resulted in an improvement of the model fit, with a decrease of chi-square and a substantial improvement of the other FIs (Jreskog & Srbom, 1989). Elimination was performed one at a time starting with the one with the highest MI and so on, since changing a single parameter in a model could affect other parts of the solution (MacCallum et al., 1992). After each elimination, the CFA was repeated on the elements that remained, evaluating the FI. These steps were repeated progressively until these indices indicated a good fit of the model according to the criteria specified in the previous paragraph for both the experienced and perpetrated violence scale. A balance between the empirical bases/theoretical rationale of the original scale (MSDV) and the revised bibliography was taken into account throughout the process.

Once the items that were retained in the short version of the scale were determined, reliability and other validity properties were evaluated to determine whether, in addition to their dimensionality, they also had good psychometric properties.

The convergent validity of the two subscales (victimization and perpetration) of the MSDV 2.0 and their dimensions were tested with the DASS-21, for its three subscales (Depression, anxiety and stress) and the self-perception of health item. We anticipated that they would correlate based on previous literature (Comecanha & Maia, 2018; Rothman et al., 2021; Wong et al., 2018). The starting hypothesis was that both subscales would present a strong and positive relationship. Bivariate correlations were performed with Spearman's rho coefficient because the Kolmogorov–Smirnov test showed no normality in quantitative variables. The interpretation of

the values was made according to Hernández et al. (2010), with the following correlation ranges: from .91 to 1.00 perfect, .76 to .90 very strong, .51 to .75 considerable, .11 to .50 medium, .01 to .10 weak and .00 no correlation.

2.7.2 | Reliability analysis

Internal consistency was determined by Cronbach's alpha test, with an acceptable value of >.7 (Nunnally & Bernstein, 1994).

2.7.3 | Cut-off points and receiver operating characteristic curves

Cut-off points were established for each dimension. Receiver operating characteristic (ROC) curves were used with the proposed criterion variables. Youden's *J* statistic was calculated for all the points on the ROC curve, and the maximum value of the index was used as a criterion for selecting the optimal cut-off point to obtain better sensitivity and specificity (Schisterman et al., 2005), with an area under the curve >0.7 (Swets, 1988).

3 | RESULTS

3.1 | Sample characteristics

Participants were 85% women and 15% men, with a mean age of 20.1 years (*SD* 1.67). Ninety-six percent (*n* = 1051) were of Spanish nationality and 4% (*n* = 40) of other nationalities (Italian, Brazilian, Moroccan and Portuguese); 91% lived in urban areas. The mean length of time in a dating relationship was 23.6 months (*SD* 19.31), where 81.7% were in a dating relationship at the time of questionnaire administration and only 5% were living together.

3.2 | Validity analysis

3.2.1 | Confirmatory factor analysis

The CFA of the proposed model, with 42 items per subscale, did not obtain good adjustment indices. Items were progressively eliminated until an adequate fit of the data to the model was achieved, obtaining 18 items that fit well to a five-factor model for each of the subscales (Table 1).

The fit indices were excellent for both the victimization subscale chi-squared (χ^2 = 503.168 (*p* <.0001); CFI = 0.946; TLI = 0.934; RMSEA = 0.053 (90% CI: 0.048–0.058) and the perpetration subscale (χ^2 = 509.859 (*p* <.0001); CFI = 0.926; TLI = 0.909; RMSEA = 0.053 (90% CI: 0.048–0.058). Figure 2a,b shows the final models selected with their factorial loads. As can be seen, in the victimization subscale all items have loads >0.53, except item 7 which

TABLE 1 Dimensions and items of the MSDV 2.0

Dimensions	Items	Victimization subscale					Perpetration subscale				
		M	SD	% Floor effect	% Ceiling effect	Factorial loads	M	SD	% Floor effect	% Ceiling effect	Factorial loads
Cyber bullying	1. Send WhatsApp messages insistently, or other types of messages through social networks	2.48	1.18	21.7	7.3	0.75	1.98	0.94	33.5	2.4	0.57
	2. Spy on each other's activity on social media: comments on photos uploaded by friends to know what they say, what they do and with whom	2.13	1.14	36.9	4.3	0.78	1.91	0.95	38.6	2.2	0.71
	3. Monitor the time of the last connection in WhatsApp messages and/or the other person's social networks	2.16	1.19	36.7	6.4	0.80	1.88	0.97	40.6	2.7	0.64
Control and surveillance	1. Make unsolicited gifts or favours	2.55	1.22	23.2	8.8	0.17	2.64	1.24	21.5	9.6	0.26
	2. Deliberately frequent the places where the other person is usually to be found (home, work, bars, party, etc.)	1.74	0.99	54.1	2.0	0.53	1.59	0.86	59.4	1.1	0.40
	3. Ask where the other person is 'every minute of the day' and/or what the other person is doing	1.56	0.98	67.8	3.3	0.74	1.41	0.68	68.0	0.5	0.51
	4. Try to make the other person feel guilty for not spending enough time together	1.78	1.08	55.7	3.8	0.74	1.52	0.72	59.6	0.1	0.60
	5. Check up with friends, family or other means, whether it is true that the other person was where s/he was	1.38	0.78	75.9	0.5	0.66	1.49	0.74	63.2	0.1	0.61
Psycho-emotional	1. Bring up something from the past to do harm	1.87	1.06	46.7	3.8	0.80	1.62	0.81	54.1	0.9	0.74
	2. Blame him/her for things that do not go well	1.66	0.96	58.0	2.4	0.77	1.44	0.69	65.0	0.4	0.66
	3. Avoid or refuse to talk to the other person (for a long time) when the perpetrator is angry	1.90	1.07	46.1	3.8	0.83	1.71	0.92	52.2	2.0	0.53
Physical	1. Physically harm someone known to either the victim	1.07	0.34	95.4	0.1	0.61	1.02	0.17	98.3	0.0	0.74
	2. Physically assault the other person in a serious way (slap, punch, etc.)	1.09	0.41	93.5	0.3	0.76	1.02	0.18	97.9	0.0	0.89

TABLE 1 (Continued)

Dimensions	Items	Victimization subscale					Perpetration subscale				
		M	SD	% Floor effect	% Celling effect	Factorial loads	M	SD	% Floor effect	% Celling effect	Factorial loads
Sexual	1. Not request consent to sexual intercourse	1.24	0.66	83.7	1.0	0.76	1.07	0.37	95.7	0.2	0.41
	2. Take advantage of the other being drunk or drugged to have sex	1.13	0.42	89.6	0.1	0.53	1.03	0.21	97.7	0.0	0.83
	3. Ask for some sex act that the other person does not wish to perform, such as using dangerous objects, or having unwanted sex with other people	1.22	0.58	84.3	0.5	0.61	1.04	0.22	96.7	0.0	0.75
	4. Push for sexual practices without a condom	1.38	0.82	76.9	1.9	0.59	1.05	0.29	95.8	0.0	0.57
	5. Touching of a sexual nature without the other person's consent	1.25	0.63	82.0	0.7	0.78	1.05	0.31	96.4	0.1	0.54

Abbreviations: factorial loads, factor loadings of the item in its dimension; M, mean; SD, standard deviation.

has a load lower than the value determined as adequate (0.17). In the perpetration subscale all the items had loads >0.40, except item 7 again, which, as in the other subscale, presents a lower load (0.26).

3.2.2 | Convergent validity

The victimization subscale score showed significant positive correlations in medium magnitude range with the self-perceived health item (rho = 0.380, $p < .01$), where higher levels of violence experienced correlated with worse levels of self-perceived health. It also correlated with the DASS-21 score (rho = 0.418, $p < .01$), showing higher levels of depression (rho = 0.444, $p < .01$), anxiety (rho = 0.470, $p < .01$) and stress (rho = 0.182, $p < .01$) those young people who presented greater violence suffered.

In turn, the subscale perpetration score also correlated positively and significantly, but in a smaller range, with self-perception of health (rho = 0.179, $p < .01$), the DASS-21 (rho = 0.225, $p < .01$) and its three dimensions: depression (rho = 0.210, $p < .01$), anxiety (rho = 0.267, $p < .01$) and stress (rho = 0.132, $p < .01$). All correlations can be seen in Table 2.

3.3 | Reliability analysis. Internal consistency

Since the CFA showed that the MSDV 2.0 is a multidimensional scale, Cronbach's alpha was evaluated for each dimension. For victimization violence, the reliability coefficients ranged from 0.703 to 0.828. For perpetration items, the alpha values ranged from 0.703 to 0.869. The lowest values were found in the control and surveillance dimension of both subscales. In general, the subscales of the MSDV presented good internal consistency (Table 3).

3.4 | Cut-off points and receiver operating characteristic curves

A cut-off score for each of the five dimensions was identified to assess their accuracy and predictive value in young people who had been victims of or had perpetrated DV. All the dimensions had an area under the curve greater than 0.7. The sensitivity and specificity values were greater than 72% in the subscale victimization, except for the sensitivity of the physical dimension, which was 64.7%. In the perpetration subscale, the values of sensitivity and specificity and sensitivity showed a more limited balance and the cut-off points for the dimensions of cyberbullying and sexual violence were lower, compared with those obtained for the subscale of victimization (Table 4).

Also, odds ratios were performed out to find the relationship between being a victim and a perpetrator (Table A2). Finally, we compared those participants who reported yes/no to each of the variables criteria with their mean score by dimensions in the corresponding subscale, which can give us a direct validation of the scale (Table A3).
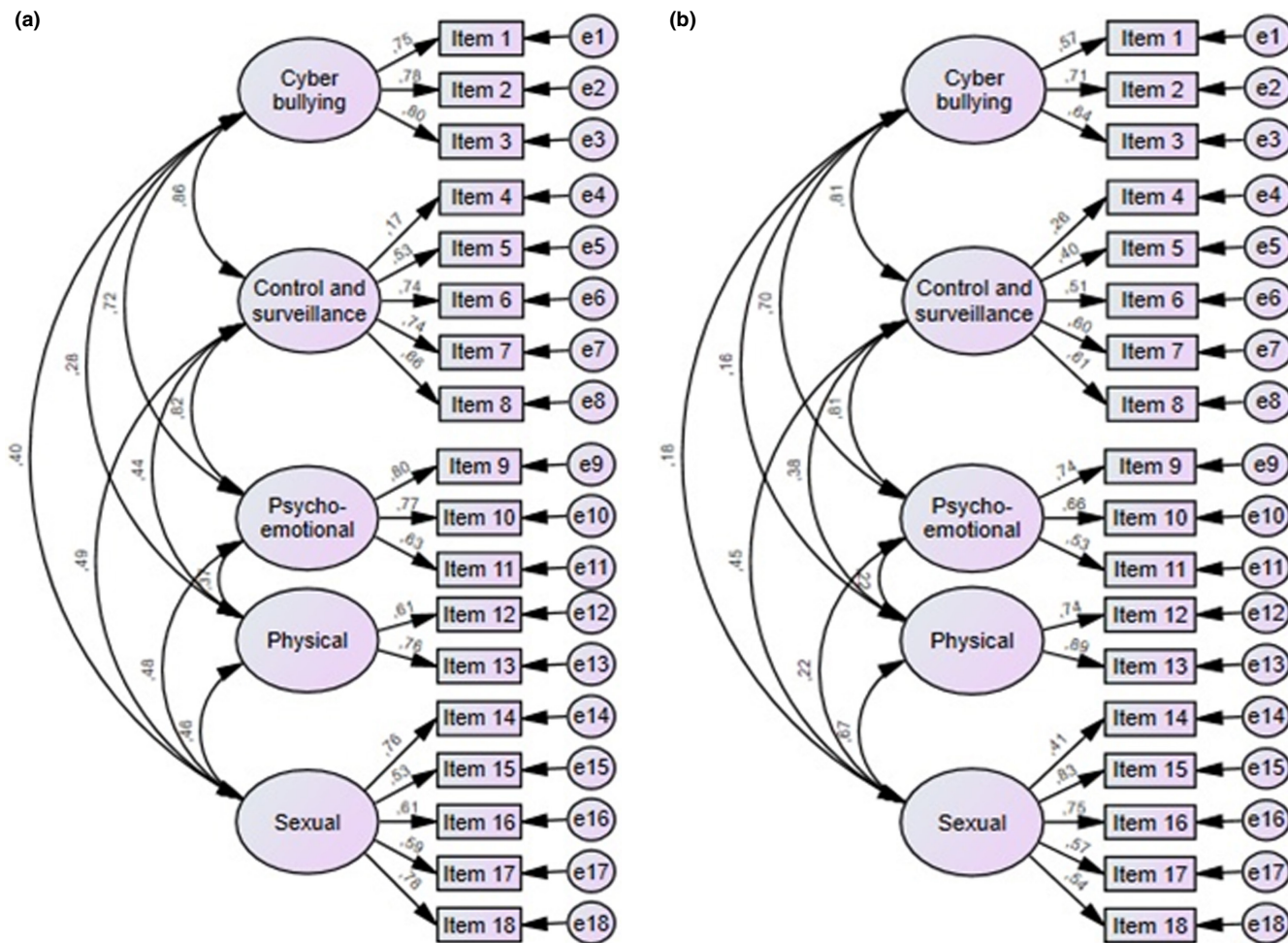
FIGURE 2 (a) CFA diagram victimisation MSDV 2.0. (b) CFA diagram perpetration MSDV 2.0

3.5 | Differences according to sex

In the violence victimization subscale, statistically significant differences were found according to sex ($p < .05$), with females scoring significantly higher as victims in the cyber bullying, control and surveillance and sexual dimensions, and males as aggressors in the psycho-emotional, physical and sexual dimensions with statistically significant differences.

However, the statistically significant differences found should be taken with caution, as they show a small effect size (PSest ≤0.56; Table 5).

3.6 | Evaluation of the quality of MSVN 2.0

The quality of the instrument was independently assessed by two psychometric experts using the COSMIN Risk of Bias checklist (Mokkink et al., 2019). Scores were obtained between Very Good to Adequate in the properties that were developed: PROM development; Content validity; Structural validity; Internal consistency; Criterion validity; Hypotheses testing for construct validity (Table A4).

4 | DISCUSSION

The results of this study show the development of a tool with solid psychometric properties, showing that it is a reliable and valid alternative to other widely used instruments such as the CTS-2 (Anderson & Leigh, 2011) or the CADRI-S (Fernández et al., 2012), which analyse DV in young people.

An update and reduction of its items has been achieved, resulting in a brief instrument of 18 items for each of its subscales (victimization and perpetration). It measures five dimensions of DV: cyber bullying, control and surveillance, psycho-emotional, physical and sexual. It shows high and adequate values regarding the quality of its psychometric properties (reliability and validity). In addition, as a differentiating feature to instruments available in the literature, cut-off points were calculated for each dimension and could be used as diagnostic criteria.

Compared with the 32-item MSDV, the MSDV 2.0 showed somewhat lower internal consistency values decreasing from 0.93 (MSDV) to 0.897 (MSDV 2.0) in the violence victimization subscale and from 0.91 (MSDV) to 0.802 (MSDV 2.0) in the violence perpetration subscale but achieved a more homogeneous structure for the

TABLE 2 Spearman correlation coefficients between MSDV 2.0, DASS 21 and self-perception of health

	DASS (Depression)	DASS (Anxiety)	DASS (Stress)	DASS-21	Self-perception of health
Cyber bullying (A)	.345[**]	.354[**]	.106[**]	.316[**]	.320[**]
Control and surveillance (A)	.368[**]	.398[**]	.148[**]	.351[**]	.312[**]
Psycho-emotional (A)	.367[**]	.394[**]	.147[**]	.346[**]	.306[**]
Physical (A)	.222[**]	.251[**]	.093[**]	.220[**]	.243[**]
Sexual (A)	.323[**]	.343[**]	.148[**]	.301[**]	.262[**]
Total (A)	.444[**]	.470[**]	.182[**]	.418[**]	.380[**]
Cyber bullying (B)	.153[**]	.209[**]	.105[**]	.177[**]	.150[**]
Control and surveillance (B)	.170[**]	.200[**]	.088[**]	.172[**]	.140[**]
Psycho-emotional (B)	.251[**]	.289[**]	.186[**]	.263[**]	.179[**]
Physical (B)	.095[**]	.105[**]	.083[**]	.100[**]	.085[**]
Sexual (B)	.062[*]	.123[**]	.037	.078[*]	.045
Total (B)	.210[**]	.267[**]	.132[**]	.225[**]	.179[**]

Note: (A) = Victimization subscale; (B) = Perpetration subscale; DASS-21 = Depression Anxiety Stress Scales-21.

*Correlation is significant at the .05 level (two-tailed).; **Correlation is significant at the .01 level (two-tailed).

TABLE 3 Reliability. Internal consistency of the MSDV 2.0

Dimensions	Items	Score range	Victimization subscale (α)	Perpetration subscale (α)
Cyber bullying	1–3	3–15	.818	.704
Control and surveillance	4–8	5–25	.703	.702
Psycho-emotional	9–11	3–15	.776	.720
Physical	12–13	2–10	.772	.869
Sexual	14–18	5–25	.828	.821
Total	1–18	18–90	.879	.802

Note: α: Cronbach's alpha (internal consistency).

two subscales. The same items were obtained by dimension for both violence perpetrated and suffered. In addition, the adjustment indices (CFI, TLI, and RMSEA) in the CFA were improved and a convergent validity was achieved, which had not previously been the case, thereby obtaining a more reliable tool.

In terms of its convergent validity, the victimization subscale and its five dimensions showed significant positive correlations with anxiety, stress and depression scores, suggesting that the more violence one suffers, the worse the depression, stress and anxiety states, these results proving consistent with those of other recent research (Comecanha & Maia, 2018; Rothman et al., 2021; Wong et al., 2018). Perpetration subscale also showed a significant and positive correlation, although to a lesser extent with anxiety, stress and depression scores, as in the Rothman et al. (2021) study (Rothman et al., 2021), where it is suggested that young people who exercise violence also have worse mental health, but to a lesser extent than girls. These results are consistent with our starting hypothesis.

The analysis of the cut-off points for each dimension can be considered a good diagnostic guideline. In our research, the

probability of correctly classifying a young person as experiencing or perpetrating DV is higher than 70%, as the area under the curve on all dimensions was higher than 0.7 with a 95% confidence interval. Sensitivity and specificity on the perpetration subscale were somewhat lower, indicating worse discrimination between true positives and true negatives. Also in this subscale, cut-off points were lower on two of its dimensions (cyber bullying and sexual). These values can be justified by the low rate of young people who identified themselves as perpetrators of violence. Despite this, research carried out in the field of IPV has used a direct, self-reported question to perform criterion validity or the validity of known groups (Comecanha & Maia, 2018), since it is difficult to recruit a sample that identifies itself as a victim/perpetrator (Tolman, 1999) or find a gold standard that adapts to the dimensions we wish to validate, as in our case.

The analysis of sex differences showed statistically significant differences ($p < .05$) for cyber bullying, control and surveillance, and sexual dimensions of victimization, where women scored significantly higher as victims of these behaviours, which is consistent with previous research (García-Carpintero et al., 2018; Macrosurvey of

TABLE 4 Area under curve, cut-off points, sensitivity and specificity

Dimension	CV		AUC (95% CI)	Cut-off point	Score range	S	E
	Yes (%)	No (%)					
Victimization subscale							
Cyber bullying	24.7	75.3	0.840 (95% CI = [0.813, 0.867])	8	3–15	73.2	79.7
Control and surveillance	14.3	85.7	0.815 (95% CI = [0.779, 0.850])	10	5–25	73.7	73.6
Psycho-emotional	20.8	79.2	0.813 (95% CI = [0.783, 0.844])	6	3–15	79.7	72.1
Physical	6.2	93.8	0.808 (95% CI = [0.738, 0.878])	3	2–10	64.7	96.7
Sexual	3.9	96.1	0.922 (95% CI = [0.885, 0.958])	8	5–25	81.4	85.8
Perpetration subscale							
Cyber bullying	10.6	89.4	0.825 (95% CI = [0.787, 0.863])	7	3–15	74.1	75.9
Control and surveillance	6.1	93.9	0.710 (95% CI = [0.649, 0.770])	10	5–25	62.7	68.8
Psycho-emotional	6.0	94	0.765 (95% CI = [0.706, 0.824])	6	3–15	64.6	74.6
Physical	1.3	98.7	0.705 (95% CI = [0.533, 0.877])	3	2–10	42.9	98.0
Sexual	0.2	99.8	0.729 (95% CI = [0.273, 1.0])	6	5–25	50.0	91.9

Abbreviations: AUC (95% CI), area under the curve (95% confidence interval); CV, criteria variable (yes/no, percentage of the total sample answered by 'yes/no' to the proposed question for the criteria variable; cut-off point, calculated by Youden's J statistic; E, specificity; S, sensitivity.

Violence against Women, 2019). In relation to psycho-emotional violence, this is the one with the highest prevalence rate among young couples, as stated in the latest Macrosurvey of Violence against Women 2019, carried out in Spain, which found that the highest proportion of controlling psychological violence is among those between 18 and 24 years old (44.3% of those who have or have had a partner).

On the other hand, in the perpetration subscale, statistically significant sex differences were found in the psycho-emotional, physical and sexual dimensions where such acts are mostly perpetrated by men, this also coincides with other recently published research (Exner-Cortens et al., 2021; García-Carpintero et al., 2018; Taylor & Xia, 2022).

It is important to mention that scores on the two subscales have been low, which coincides with most scales that measure IPV (Anderson & Leigh, 2011; Aizpitarte et al., 2017; Benítez et al., 2014; Fernández et al., 2012; Lara et al., 2021; López et al., 2016; Rodríguez-Díaz et al., 2017; Ronzón et al., 2019; Rothman et al., 2021; Soriano et al., 2021). This may be because much of the behaviour measured can be justified in the construct and myths of romantic love and may not identified as violent (Estébanez, 2018; Rodríguez et al., 2018). It should also be considered that the research was carried out in an academic context, which promotes fair and egalitarian values. This may have conditioned the participants to score lower to reflect their role in the university. These data are consistent with other research conducted among young people and in academic contexts (Boira et al., 2017; Soriano et al., 2021). Another phenomenon in the subscale perpetration is social desirability, where the participants score lower in behaviour which is not tolerated by society (Boira et al., 2017). This may have caused the quality of the psychometric properties in the subscale of perpetration to be lower.

Finally, the results of our research show that DV is a complex problem to measure, but a key problem to detect in the

university context, coinciding with other research (Arenas-Carbellido et al., 2020; Kelmendi & Baumgartner, 2020; Kim et al., 2019). They also show the relevance of having a valid and reliable instrument like the one designed and validated in our research, the MSVN 2.0.

4.1 | Limitations

One of the limitations of this study is that the data were self-reported through an online survey, which could have limited access to the study to those participants who did not have an internet connection or electronic device. Also, due to social desirability, the participants could systematically complete their own victimization or perpetration to a low or high degree.

Second, the scores reported by the participants were low, as found in most of the research conducted in this field. To solve this, a CFA was carried out following the model of Robust maximum-likelihood estimation to mitigate the possible biases that could occur in the estimates due to the observed floor effects.

Third, the sample, although large, was a convenience sample, not representative of the university context of young Andalusians, and just 15% of participants were male. In addition, the percentage of participants of other nationalities or cultures and from rural areas was so small that these data could not be analysed for statistical inference.

Fourth, the study was validated without being able to demonstrate whether these measures are culturally appropriate for non-heteronormative identities.

Finally, in relation to the quality review of the instrument by psychometric experts, at least a test–retest reliability and criterion validity with a gold standard should have been performed. In spite of this, sufficient psychometric tests were performed with high–adequate quality. The investigators undertake to consider all these

TABLE 5 Differences according to sex

Dimensions	Sex (N)	Mean ranks	p	PSest
Cyber bullying (A)	Women (926)	560.04		
	Mens (165)	467.18		
	Total (1091)		.000*	.41
Control and surveillance (A)	Women (926)	557.46		
	Mens (165)	481.69		
	Total (1091)		.004*	.43
Psycho-emotional (A)	Women (926)	552.36		
	Mens (165)	510.28		
	Total (1091)		.108*	
Physical (A)	Women (926)	544.18		
	Mens (165)	556.22		
	Total (1091)		.311	
Sexual (A)	Women (926)	558.97		
	Mens (165)	473.20		
	Total (1091)		.000*	.42
Cyber bullying (B)	Women (926)	549.46		
	Mens (165)	526.61		
	Total (1091)		.385	
Control and surveillance (B)	Women (926)	550.80		
	Mens (165)	519.06		
	Total (1091)		.230	
Psycho-emotional (B)	Women (926)	536.59		
	Mens (165)	598.79		
	Total (1091)		.016*	.44
Physical (B)	Women (926)	541.95		
	Mens (165)	568.74		
	Total (1091)		.000*	.48
Sexual (B)	Women (926)	533.54		
	Mens (165)	615.95		
	Total (1091)		.000*	.42

Note: (A), Dimensions of the Victimization subscale; (B), Dimensions of the Perpetration subscale; N, sample size; p, p value (*$p < .05$); PSest, probability of superiority (effect size).

limitations in future research, as well as, the opportunity of validating this instrument in other community and clinical contexts.

4.2 | Implications for research, policy and practice

A reliable and valid instrument has been built to objectively measure the presence of DV. In addition, this instrument, being an abbreviated version, can be more rapidly completed in various environments where due to lack of time a quick analysis of the situation is necessary, such as in the university and clinical care.

MSDV 2.0 could be used in future studies to analyse the most prevalent violent behaviours among young university students and probe the causes and consequences. These data can be used to design or modify protocols against DV and to elaborate guidelines for healthy attitudes in relationships. All this will provide knowledge for designing future policies at the local, regional and national levels to help eradicate this violence.

5 | CONCLUSION

This study has made it possible to obtain the MSDV 2.0, an updated and shorter version than the original scale (MSDV) but incorporating new emerging and prevalent behaviours among young university students. In addition, cut-off points have been identified for each subscale that will discriminate the presence of different types of DV.

It is important to have an instrument such as the one designed that has proven to be valid and reliable in the university context, since institutions of higher education have a duty to educate, train and inculcate the values of a modern, democratic, egalitarian, free and supportive society.

In addition, it is necessary to monitor these behaviours, present figures and, based on them, design prevention programs and policies, where nurses have a key role in both the educational context and clinical care. This, is all the more important among young people, since these behaviours can be the prelude to consolidating and normalizing violent relationships in future relationships and in adult life.

Finally, this research agrees with the Sustainable Development Goal 5 (SDG5), to achieve gender equality and to empower women and girls, calls for an end to all violence against women, attributing most of this violence to intimate partners.

AUTHOR CONTRIBUTIONS

ACKNOWLEDGEMENTS

CONFLICT OF INTERESTS

No conflict of interest has been declared by the author(s).

PEER REVIEW

The peer review history for this article is available at https://publons.com/publon/10.1111/jan.15300.

DATA AVAILABILITY STATEMENT

Data available on request from the authors.

ORCID

María de los Ángeles García-Carpintero-Muñoz https://orcid.org/0000-0003-4961-484X
Lorena Tarriño-Concejero https://orcid.org/0000-0003-0158-5744
Eugenia Gil-García https://orcid.org/0000-0003-3862-6845
Ana María Pórcel-Gálvez https://orcid.org/0000-0001-8461-6379
Sergio Barrientos-Trigo https://orcid.org/0000-0001-7618-3091

REFERENCES

Abad, F., Olea, J., Ponsoda, V., & García, C. (2011). *Measurement in social and health sciences*. Synthesis.

Aizpitarte, A., Alonso, I., de Vijver, F. J. R., Perdomo, M. C., Galvez-Sobral, J. A., & Garcia-Lopez, E. (2017). Development of a dating violence assessment tool for late adolescence across three countries: The Violence in Adolescents' Dating Relationships Inventory (VADRI). *Journal of Interpersonal Violence*, 32(17), 2626–2646. https://doi.org/10.1177/0886260515593543

Anderson, M. L., & Leigh, I. W. (2011). Intimate partner violence against deaf female college students. *Violence Against Women*, 17(7), 822–834. https://doi.org/10.1177/1077801211412544

Arbuckle, J. L. (2014). *Amos* (Version 23.0) [Computer Program]. IBM SPSS.

Arenas-Carbellido, C., Arias, A., & Emery, C. R. (2020). Dating violence heterogeneity in University Couples. The Cases of Spain and South Korea. *Social Work in Public Health*, 35(6), 368–379. https://doi.org/10.1080/19371918.2020.1775747

Bagwell, M. E. (2021). Women's experiences of sexual violence in intimate relationships: Applying a new taxonomy. *Journal of Interpersonal Violence*, 36(13–14), NP7813–NP7839. https://doi.org/10.1177/0886260519827667

Banyard, V. L., & Cross, C. (2008). Consequences of teen dating violence: Understanding intervening variables in ecological context. *Violence Against Women*, 14(9), 998–1013. https://doi.org/10.1177/1077801208322058

Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin*, 107(2), 238–246. https://doi.org/10.1037/0033-2909.107.2.238

Boira, S., Chilet-Rosell, E., Jaramillo-Quiroz, S., & Reinoso, J. (2017). Sexism, distorted thoughts and violence in relationships in university students from Ecuador in areas related to well-being and health. *Universitas Psychologica*, 16, 1–12. https://doi.org/10.11144/javeriana.upsy16-4.spdv

Bosch, E., Herrezuelo, R., & Ferrer, V. A. (2019). Romantic love, as renunciation and sacrifice: What do young people think? *FEMERIS: Multidisciplinary Journal of Gender Studies*, 4(3), 184–202.

Brown, T.A. (2015). *Confirmatory factor analysis for applied research*. Guilford. 2 ed. ISBN: 9781462515363

Cava, M. J., Tomás, I., Buelga, S., & Carrascosa, L. (2020). Loneliness, depressive mood and cyberbullying victimization in adolescent victims of cyber dating violence. *International Journal of Environmental Research and Public Health*, 15(12), 4269. https://doi.org/10.3390/ijerph17124269

Cha, S., Ihongbe, T. I., & Masho, S. W. (2016). Racial and gender differences in dating violence victimization and disordered eating among U.S. high schools. *Journal of Women's Health*, 25(8), 791–800. https://doi.org/10.1089/JWH.2015.5324

Comecanha, R., & Maia, A. (2018). Screening tool for psychological intimate partner violence: Portuguese validation of the psychological maltreatment of women inventory. *Violence and Victims*, 33(1), 75–90. http://doi.org/10.1891/0886-6708.VV-D-16-00060

IBM Corp. (2019). *IBM SPSS statistics for windows*, Version 26.0. IBM Corp.

Cruz, Ó. A., & Muñoz, A. I. (2015). Validation of an instrument to identify the level of vulnerability of health workers to tuberculosis in health institutions (IVTS TB-001). *Occupational Medicine and Safety*, 61(241), 448–467.

Davila, J., Capaldi, D. M., & Greca, A. M. L. (2016). Adolescent/young adult romantic relationships and psychopathology. *Development and Psychopathology*, 1, 1–34. https://doi.org/10.1002/9781119125556.DEVPSY114

Demissie, Z., Clayton, H., Vivolo-Kantor, A., & Estefan, L. (2018). Sexual teen dating violence victimization: Associations with sexual risk behaviors among US high school students. *Violence & Victimization*, 33(5), 964–980. https://doi.org/10.1891/0886-6708.33.5.964

Emelianchik, K., Hays, D. G., & Hill, T. (2018). Initial development of the teen screen for dating violence: Exploratory factor analysis, rasch model, and psychometric data. *Measurement and Evaluation in Counseling and Development*, 51(1), 16–31. https://doi.org/10.1080/07481756.2017.1395707

Estébanez, I. (2018). *Cyberviolence towards teenage girls on social networks*. Andalusian Institute of Women. Ministry of Equality and Social Policies. https://www.juntadeandalucia.es/export/drupaljda/guia-ciberviolencia-adolescentes.pdf

European Union Agency for Fundamental Rights. (2014). *Violence against women: An EU-wide survey. Results at a glance*. FRA – European Union Agency for Fundamental Rights. https://fra.europa.eu/en/publication/2014/violence-against-women-eu-wide-survey-results-glance

Exner-Cortens, D., Baker, E., & Craig, W. (2021). The national prevalence of adolescent dating violence in Canada. *The Journal of Adolescent Health*, *69*(3), 495–502. https://doi.org/10.1016/j.jadohealth.2021.01.032

Fernández, L., Wekerle, C., & Goldstein, A. L. (2012). Measuring adolescent dating violence: Development of "conflict in adolescent dating relationships inventory" short form. *Advances in Mental Health*, *11*(1), 35–54. https://doi.org/10.5172/jamh.2012.11.1.35

Fonseca, E., Paino, M., Lemos, S., & Muñiz, J. (2010). Psychometric properties of the Depression Anxiety and Stress Scales-21 (Dass-21) in Spanish college students. *Ansiedad Estrés*, *16*(2/3), 215–226.

Gabelas, J., & Marta, C. (2020). *The TRIC era: R-elational factor and edu-communication*. Editions Egregius.

García-Carpintero, M. Á., Rodríguez, J., & Porcel, A. M. (2018). Design and validation of the scale for the detection of dating violence in young people at the University of Seville. *Gaceta Sanitaria*, *32*(2), 121–128. https://doi.org/10.1016/j.gaceta.2017.09.006

Government Delegation against Gender Violence, Spain. (2019). *Macrosurvey on violence against women 2019*. https://violenciagenero.igualdad.gob.es/violenciaEnCifras/macroencuesta2015/pdf/Macroencuesta_2019_estudio_investigacion.pdf

Grissom, R. (1994). Probability of the superior outcome of one treatment over another. *Journal of Applied Psychology*, *79*(2), 314–316. https://jglobal.jst.go.jp/en/detail?JGLOBAL_ID=200902161044172246

Hernández, R., Fernández, C., & Baptista Lucio, P. (2010). *Research methodology*. Mcgraw-Hill & I. México.

Hu, L., & Bentler, P. M. (2009). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, *6*(1), 1–55. https://doi.org/10.1080/10705519909540118

Idler, E. L., & Benyamini, Y. (1997). Self-rated health and mortality: A review of twenty-seven community studies. *Journal of Health and Social Behavior*, *38*(1), 21–37. https://doi.org/10.2307/2955359

Integrated University Information System. (2021). Data and numbers of the Spanish University System (Publication 2020–2021). https://www.universidades.gob.es/stfls/universidades/Estadisticas/ficheros/Datos_y_Cifras_2020-21.pdf

Jack, S. M., Kimber, M., Davidov, D., Ford- Gilboe, M., Wathen, C. N., McKee, C., Tanaka, M., Boyle, M., Johnston, C., Coben, J., Gasbarro, M., McNaughton, D., O'Brien, R., Olds, D. L., Scribano, P., & MacMillan, H. L. (2021). Nurse-family partnership nurses' attitudes and confidence in identifying and responding to intimate partner violence: An explanatory sequential mixed methods evaluation. *Journal of Advanced Nursing*, *77*, 3894–3910. https://doi.org/10.1111/jan.14979

Jreskog, K. G. & Srbom, D. (1989). LISREL 7 A guide to the program and applications. (I. SPSS [ed.]).

Kelmendi, K., & Baumgartner, F. (2020). Exploring violence socialization and approval of intimate partner violence among University Students in Kosovo. *Journal Interpersonal Violence*, *35*(5–6), 1081–1107. https://doi.org/10.1177/0886260517692336

Khanhkham, A., Williams, R. D., & Housman, J. M. (2020). Sexual dating violence, school-based violence, and risky behaviors among U.S. High School Students. *Journal of Community Health*, *45*, 932–942. https://doi.org/10.1007/s10900-020-00811-1

Kim, Y., Lee, E., & Lee, H. (2019). Sexual double standard, dating violence recognition, and sexual assertiveness among university students in South Korea. *Asian Nursing Research*, *13*(1), 47–52. https://doi.org/10.1016/j.anr.2019.01.003

Kisa, S., & Zeyneloğlu, S. (2019). Perceptions and predictors of dating violence among nursing and midwifery students. *Journal of Advance Nursing*, *75*, 2099–2109. https://doi.org/10.1111/jan.13982

Llano-Suárez, A., Lana, A., Gasch-Gallén, Á., & Fernández-Feito, A. (2021). Gender roles and intimate partner violence among female university students in Spain: A cross-sectional study. *PLoS One*, *16*(11), e0259839. https://doi.org/10.1371/JOURNAL.PONE.0259839

López, J., Fabelo, H., Rodríguez, L., & Rodríguez, F. (2016). The dating violence questionnaire: Validation of the Cuestionario de Violencia de Novios using a college sample from the United States. *Violence and Victims*, *31*(3), 438–456. https://doi.org/10.1891/0886-6708.VV-D-14-00077

MacCallum, R. C., Roznowski, M., & Necowitz, L. B. (1992). Model modifications in covariance structure analysis: The problem of capitalization on chance. *Psychological Bulletin*, *111*(3), 490–504. https://doi.org/10.1037/0033-2909.111.3.490

Mokkink, L., Princen, C. A., Patrick, D., Alonso, J., Bouter, L. M., de Vet, H., & Terwee, C. B. (2019). *COSMIN Study Design checklist for Patient-reported outcome measurement instruments*. Department of Epidemiology and Biostatistics Amsterdam Public Health Research Institute Amsterdam University Medical Centers. https://www.cosmin.nl/wp-content/uploads/COSMIN-study-designing-checklist_final.pdf

Morales, P., Urosa, B., & Blanc, A. (2003). *Construction of Attitudes Likert Scales*. Editorial La Muralla, Vol. 26.

Moreno, C., Ramos, P., Rivera, F., Jiménez, A., & García, I. (2012). Behaviors related to the health and development of Spanish adolescents. Study Summary Health Behaviour in School Aged Children (HBSC-2010). https://idus.us.es/bitstream/handle/11441/34137/Las%20conductas%20relacionadas%20con%20la%20salud.%20Adolescentes%20espa%c3%b1oles.pdf?sequence=1&isAllowed=y

Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory*. McGraw-Hill, Inc.

Organic Law 3/2018, of December 5, on the Protection of Personal Data and Guarantee of Digital Rights. (2018). Head of state. https://www.boe.es/buscar/doc.php?id=BOE-A-2018-16673

Penfield, R. D., & Giacobbi, P. R. (2004). Applying a score confidence interval to Aiken's item content-relevance index. *Measurement in Physical Education and Exercise Science*, *8*(4), 213–225. https://doi.org/10.1207/s15327841mpee0804_3

Pérez-Marco, A., Soares, P., Davó-Blanes, M. C., & Vives-Cases, C. (2020). Identifying types of dating violence and protective factors among adolescents in Spain: A qualitative analysis of lights4violence materials. *International Journal of Environmental Research and Public Health*, *17*(7), 2443. https://doi.org/10.3390/IJERPH17072443

Perles, F., San Martín, J., & Canto, J. (2019). Gender and conflict resolution strategies in Spanish teen couples: Their relationship with jealousy and emotional dependency. *Journal of Interpersonal Violence*, *34*(7), 1461–1486. https://doi.org/10.1177/0886260516651316

Rodríguez, C., Durán, M., & Martínez, R. (2018). Cyber aggressors in dating relationships and its relation with psychological violence, sexism and jealousy. *Health and Addictions*, *18*(1), 17–27.

Rodríguez-Díaz, J. F., Herrero, J., Rodriguez-Franco, L., Bringas-Molleda, C., Paino-Quesada, S. G., & Perez, B. (2017). Validation of Dating Violence Questionnarie-R (DVQ-R). *International Journal of Clinical and Health Psychology*, *17*(1), 77–84. https://doi.org/10.1016/j.ijchp.2016.09.001

Rodríguez-Santero, J., García-Carpintero, M. Á., & Porcel, A. M. (2017). Love attitudes styles amongst college students. Differences by sex-gender. *Revista Internacional de Sociología*, *75*(3), e073. https://doi.org/10.3989/ris.2017.75.3.15.171

Ronzón, R. C., Muñoz-Rivas, M. J., Zamarrón Cassinello, M. D., & Redondo Rodríguez, N. (2019). Cultural Adaptation of the Modified Version of the Conflicts Tactics Scale (M-CTS) in Mexican Adolescents. *Frontiers in Psychology*, *10*, 619. https://doi.org/10.3389/fpsyg.2019.00619

Rothman, E. F., Cuevas, C. A., Mumford, E. A., Bahrami, E., & Taylor, B. G. (2021). The psychometric properties of the measure of adolescent relationship harassment and abuse (MARSHA) with a nationally representative sample of U.S. Youth. *Journal of Interpersonal Violence*, 088626052098548. https://doi.org/10.1177/0886260520985480

Rubio, F., Carrasco, M. Á., Amor, P. J., & López, M. A. (2015). Adolescent dating violence related factors: A critical review. *Anuario de*

Psicología Jurídica, 25(1), 47–56. https://doi.org/10.1016/j.apj.2015.01.001

Schisterman, E. F., Perkins, N. J., Liu, A., & Bondell, H. (2005). Optimal cut-point and its corresponding Youden Index to discriminate individuals using pooled blood samples. *Epidemiology*, 16(1), 73–81. https://doi.org/10.1097/01.ede.0000147512.81966.ba

Sörbom, D. (1989). Model modification. *Psychometrika*, 54(3), 371–384.

Soriano, E., Sanabria, M., & Cala, V. C. (2021). Design and validation of the scale tdv-vp teen dating violence: Victimisation and perpetration for Spanish speakers. *International Journal of Environmental Research and Public Health*, 18(2), 1–15. https://doi.org/10.3390/ijerph18020421

Spanish Network of Healthy Universities. (2021). *Spanish Network of Healthy Universities: The University as a health-promoting environment*. https://www.sanidad.gob.es/en/profesionales/saludPublica/prevPromocion/promocion/UniversidadesSaludables/REUS.htm

Straus, M. A., & Gozjolko, K. L. (2014). "Intimate Terrorism" and gender differences in injury of dating partners by male and female university students. *Journal of Family Violence*, 29(1), 51–65. https://doi.org/10.1007/s10896-013-9560-7

Swets, J. (1988). Measuring the accuracy of diagnostic systems. *Science*, 240(4857), 1285–1293.

Taylor, S., & Xia, Y. (2022). Dating violence among rural adolescents: Perpetration and victimization by gender. *Journal of Interpersonal Violence*, 37(9-10), 886260520971613. https://doi.org/10.1177/0886260520971613

Tolman, R. M. (1999). The validation of the psychological maltreatment of women inventory. *Violence and Victims*, 14, 25–37.

Tucker, L. R., & Lewis, C. (1973). A reliability coefficient for maximum likelihood factor analysis. *Psychometrika*, 38(1), 1–10.

United Nations Entity for Gender Equality and the Empowerment of Women (UN Women). (2021). *Facts and figures: Ending violence against women*. https://www.unwomen.org/en/what-we-do/ending-violence-against-women/facts-and-figures

Van Teijlingen, E. R., & Hundley, V. (2002). The importance of pilot studies. *Nursing Standard*, 16(40), 33–36. https://doi.org/10.7748/ns2002.06.16.40.33.c3214

Wang, J., Hefetz, A., & Liberman, G. (2017). Applying structural equation modelling in educational research. *Cultura y Educacion*, 29(3), 563–618. https://doi.org/10.1080/11356405.2017.1367907

Wong, J. Y., Fong, D. Y., Yau, J. H., Choi, E. P., Choi, A. W., & Brown, J. B. (2018). Using the woman abuse screening tool to screen for and assess dating violence in college students. *Violence Against Women*, 24(9), 1039–1051. https://doi.org/10.1177/1077801217731542

World Health Organization (WHO). (2021). *Violence against women. Key facts*. https://www.who.int/news-room/fact-sheets/detail/violence-against-women

APPENDIX A

TABLE A1 Delphi panel analysis for the criteria of clarity, coherence and relevance of the proposed items

Items proposed in the Delphi panel for experts	Clarity			Coherence			Relevance		
	X	SD	V Aiken	X	SD	V Aiken	X	SD	V Aiken
1ª Round									
1. Send WhatsApp messages insistently, or other types of messages through social networks (R)	3.6	0.75	0.87	3.44	0.70	0.81	3.48	0.75	0.83
2. Spy on each other's activity on social media: comments on photos uploaded by friends to know what they say, what they do and with whom (R)	3.76	0.59	0.92	3.56	0.57	0.85	3.64	0.56	0.88
3. Monitor the time of the last connection in WhatsApp messages and/or the other person's social networks(R)	3.72	0.53	0.91	3.56	0.57	0.85	3.72	0.53	0.91
4. Monitor the other person's location by electronic devices (N)	3.72	0.66	0.91	3.52	0.75	0.84	3.56	0.75	0.85
5. Demand that the other person delete videos or photos from social media profiles because they find them intolerable (N)	3.68	0.61	0.89	3.64	0.62	0.88	3.68	0.61	0.89
6. Upload photos of the other person to social networks to harm and/or insult him/her on social networks (R)	3.68	0.54	0.89	3.56	0.70	0.85	3.56	0.70	0.85
7. Make unsolicited gifts or favours (O)	3.72	0.66	0.91	3.64	0.69	0.88	3.72	0.66	0.91
8. Change classes, practical sessions or the like to be closer (to him or her) (OR)	3.64	0.56	0.88	3.56	0.57	0.85	3.6	0.57	0.87
9. Wait for the other person outside class, home or work (O)	3.72	0.66	0.91	3.6	0.63	0.87	3.64	0.69	0.88
10. Deliberately frequent the places where the other person is usually to be found (home, work, bars, party, etc.) (O)	3.6	0.69	0.87	3.44	0.75	0.81	3.6	0.69	0.87
11. Stalk him/her (O)	3.72	0.53	0.91	3.44	0.75	0.81	3.44	0.75	0.81
12. Ask where the other person is 'every minute of the day' and/or what the other person is doing (R)	3.76	0.58	0.92	3.44	0.70	0.81	3.6	0.69	0.87
13. Try to get the other person to leave their hobbies so they can spend more time together (R)	3.68	0.61	0.89	3.52	0.70	0.84	3.56	0.70	0.85
14. Try to make the other person feel guilty for not spending enough time together (O)	3.56	0.57	0.85	3.48	0.64	0.83	3.52	0.64	0.84
15. Check up with friends, family or other means, whether it is true that the other person was where s/he said s/he was (O)	3.72	0.45	0.91	3.8	0.40	0.93	3.72	0.45	0.91
16. Tell the other person how they should dress, comb their hair, put on makeup, tattoo or other items of their body image (N)	3.68	0.61	0.89	3.52	0.70	0.84	3.56	0.70	0.85
17. Tell the other person that s/he is worthless, that s/he is ugly, a failure, or similar (R)	3.84	0.46	0.95	3.64	0.69	0.88	3.64	0.69	0.88
18. Ridicule the other person's physical appearance or something similar (O)	3.68	0.55	0.89	3.64	0.56	0.88	3.6	0.63	0.87
19. Insult or belittle the other person in front of others (O)	3.44	0.75	0.81	3.4	0.75	0.80	3.4	0.75	0.80
20. Threatening to leave him/her and start dating someone else who would be a better partner (R)	3.76	0.43	0.92	3.64	0.48	0.88	3.76	0.43	0.92
21. Tell him/her that someone else would be a better partner (O)	3.72	0.53	0.91	3.6	0.57	0.87	3.68	0.55	0.89
22. Bring up something from the past to do harm (O)	3.76	0.51	0.92	3.56	0.64	0.85	3.64	0.62	0.88
23. Blame him/her for things that do not go well (O)	3.84	0.46	0.95	3.68	0.61	0.89	3.72	0.60	0.91

TABLE A1 (Continued)

Items proposed in the Delphi panel for experts	Clarity			Coherence			Relevance		
	X	SD	V Aiken	X	SD	V Aiken	X	SD	V Aiken
24. Refuse to deal with a problem or talk to the other person for a long time, when the perpetrator is angry (R)	2.76	0.81	0.58	3.24	0.59	0.75	3.12	0.71	0.71
25. Become so aggressive as to scare the other person (punching or kicking the wall, a car, or hitting something) (R)	3.84	0.46	0.95	3.84	0.37	0.95	3.72	0.53	0.91
26. Threaten to physically assault a known person (O)	3.52	0.70	0.84	3.44	0.64	0.81	3.48	0.70	0.83
27. Physically harm someone known to either the victim (O)	3.52	0.70	0.84	3.44	0.64	0.81	3.48	0.70	0.83
28. Drive recklessly to scare the other person (O)	3.48	0.70	0.83	3.2	0.85	0.73	3.48	0.70	0.83
29. Threaten to break, steal or throw away the other person's personal items (O)	3.6	0.69	0.87	3.52	0.70	0.84	3.6	0.69	0.87
30. Break or throw away the other person's personal items (R)	3.92	0.27	0.97	3.8	0.40	0.93	3.88	0.32	0.96
31. Physically assault the other person in a mild way (grab, push, etc.) (O)	3.4	0.69	0.80	3.44	0.64	0.81	3.36	0.69	0.79
32. Physically assault the other person in a serious way (slap, punch, etc.) (O)	3.88	0.32	0.96	3.64	0.48	0.88	3.8	0.40	0.93
33. Threaten self-harm (O)	3.48	0.64	0.83	3.4	0.57	0.80	3.4	0.57	0.80
34. Hurt oneself or try to pressure the other person (R)	3.88	0.32	0.96	3.72	0.45	0.91	3.84	0.37	0.95
35. Sexual assault using force or threats (R)	3.96	0.20	0.99	3.8	0.40	0.93	3.88	0.32	0.96
36. Seek consent before having sex. Make sure the other person also wants to have sex and has said so (N)	2.72	1.34	0.57	3.56	0.70	0.85	3.16	1.05	0.72
37. Take advantage of the other person while inebriated or under the influence of other drugs to have sex (N)	2.8	0.85	0.6	3.2	0.63	0.73	3.12	0.71	0.71
38. Ask for some sex act that the other person does not wish to perform, such as using dangerous objects, or having unwanted sex with other people (N)	3.68	0.61	0.89	3.6	0.57	0.87	3.64	0.62	0.88
39. Push for sexual practices without a condom (N)	3.88	0.32	0.96	3.72	0.45	0.91	3.84	0.37	0.95
40. Send sexual images or photos with obscene content (N)	3.76	0.43	0.92	3.52	0.57	0.84	3.64	0.56	0.88
41. Touching of a sexual nature without the other person's consent (N)	3.96	0.20	0.99	3.72	0.45	0.91	3.84	0.37	0.95
42. Pressure or blackmail to have sex that the other person does not consent to (N)	2.92	0.84	0.64	3.12	0.65	0.71	3.12	0.71	0.71
2ª Round									
24. Avoid or refuse to talk to the other person (for a long time) when the perpetrator is angry (R*)	3.71	0.54	0.90	3.71	0.54	0.90	3.75	0.52	0.92
36. Not request consent to sexual intercourse (R*)	3.46	0.71	0.82	3.58	0.57	0.86	3.63	0.56	0.88
37. Take advantage of the other being drunk or drugged to have sex (R*)	3.50	0.71	0.83	3.71	0.61	0.90	3.58	0.57	0.86
42. Have unwanted sex, so that the other person does not get angry or out of fear of the other person (R*)	3.46	0.71	0.82	3.54	0.58	0.85	3.50	0.50	0.83

Abbreviations: N, proposed new item; O, original item of the EMVN; R, reformulated item of the MSDV; R*, item reformulated after the second round according to the contributions of the experts; SD, standard deviation; VAkein, VAkein statistic (Clarification: Item 24, in the second round was reformulated because it was in the original scale); X, mean.

TABLE A2 Odd ratios between the criteria variables of each dimension of the victimization and perpetration subscale

Criteria variables victimization subscale	Criteria variables perpetration subscale				
	Cyber bullying Yes (N = 116; 10.6%)	Control and surveillance Yes (N = 67; 6.1%)	Psycho-emotional Yes (N = 65; 6.0%)	Physical Yes (N = 14; 1.3%)	Sexual Yes (N = 2; 0.2%)
Cyber bullying					
No (N = 822; 75.3%)	23	24	28	6	1
Yes (N = 269; 24.7%)	93	43	37	8	1
p	<.001	<.001	<.001	.005	.405
OR (95% CI)	18.4 (11.3–29.8)	6.3 (3.8–1.6)	4.5 (2.7–7.5)	4.2 (1.4–12.1)	
Control and surveillance					
NO (N = 935; 85.7%)	79	19	40	9	1
Yes (N = 156; 14.3%)	37	48	25	5	1
p	<.001	<.001	<.001	.021	.149
OR (95% CI)	3.4 (2.2–5.2)	21.43 (12.1–37.8)	4.3 (2.5- 7.3)	3.4 (1.1–10.3)	
Psycho-emotional					
No (N = 864; 79.2%)	69	28	13	6	1
Yes (N = 227; 20.8%)	47	39	52	8	1
p	<.001	<.001	<.001	.001	.309
OR (95% CI)	3.0 (2.0–4.5)	6.2 (3.7–10.3)	19.5 (10.4–36.5)	5.2 (1.8–15.2)	
Physical					
No (N = 1023; 93.8%)	104	53	49	7	1
Yes (N = 68; 6.2%)	12	14	16	7	1
p	.53	<.001	<.001	<.001	.010
OR (95% CI)		4.7 (2.5–9.0)	6.1 (3.3–11.5)	16.7 (5.7–48.9)	15.3 (0.9–246.6)
Sexual					
No (N = 1048; 96.1%)	106	54	55	12	1
Yes (N = 33; 3.9%)	10	13	10	2	1
p	.006	<.001	<.001	.045	.001
OR (95% CI)	2.7 (1.3–5.6)	7.9 (3.9–16.1)	5.5 (2.6–11.7)	4.2 (0.9–19.4)	24.9 (1.5–405.4)

Note: No, sample that answered 'no' to the criterion variable related to the dimension; Yes, sample that 'Yes' answered to the criterion variable related to dimension; *p*, *p* value ($p < .05$); OR, odds ratio (95% confidence interval).

TABLE A3 Comparison between the young people who reported yes/no in each of the criterion variables and their mean ranks in the corresponding dimension

Criteria variable	Cyber bullying		Control and surveillance		Psycho-emotional		Physical		Sexual		Total Victimization subscale	
	Mean ranks	p	Mean ranks	p	Mean ranks	p	Mean ranks	p	Mean ranks	p	Mean ranks	p
Victimization subscale												
Cyber bullying												
No (N = 822)	454.53	<.001*	468.90	<.001*	467.11	<.001*	525.60	<.001*	501.69	<.001*	451.51	<.001*
Yes (N = 269)	825.50		781.59		787.06		608.33		681.41		834.73	
Control and surveillance												
No (N = 935)	492.95	<.001*	496.90	<.001*	498.66	<.001*	529.17	<.001*	514.92	<.001*	488.87	<.001*
Yes (N = 156)	863.93		840.29		829.72		646.88		732.27		888.43	
Psycho-emotional												
No (N = 864)	486.25	<.001*	495.42	<.001*	474.84	<.001*	526.59	<.001*	498.13	<.001*	471.74	<.001*
Yes (N = 227)	773.43		738.53		816.83		619.90		728.19		828.65	
Physical												
No (N = 1023)	527.50	<.001*	528.69	<.001*	524.97	<.001*	525.05	<.001*	531.04	<.001*	522.92	<.001*
Yes (N = 68)	824.26		806.42		862.35		861.15		771.11		893.20	
Sexual												
No (N = 1048)	535.05	<.001*	534.54	<0.001*	533.04	<0.001*	539.62	<0.001*	527.87	<0.001*	529.14	<0.001*
Yes (N = 43)	812.99		825.28		861.91		701.53		987.81		956.99	

Criteria variable	Cyber bullying		Control and surveillance		Psycho-emotional		Physical		Sexual		Total perpetration subscale	
	Mean ranks	p	Mean ranks	p	Mean ranks	p	Mean ranks	p	Mean ranks	p	Mean ranks	p
Perpetration subscale												
Cyber bullying												
No (N = 975)	508.34	<.001*	522.23	<.001*	516.59	<.001*	541.52	<.001*	540.69	.001*	508.76	<.001*
Yes (N = 116)	862.50		745.82		793.22		583.66		590.66		859.01	
Control and surveillance												
No (N = 1024)	529.26	<.001*	531.96	<.001*	530.68	<.001*	543.71	.001*	542.95	.009*	526.91	<.001*
Yes (N = 67)	801.80		760.58		780.10		581.06		592.69		837.81	
Psycho-emotional												
No (N = 1026)	531.90	<.001*	534.09	<.001*	528.78	<.001*	542.08	<.001*	542.86	.006*	527.44	<.001*
Yes (N = 65)	768.55		734.01		817.83		607.92		595.63		838.98	

TABLE A3 (Continued)

Criteria variable	Cyber bullying		Control and surveillance		Psycho-emotional		Physical		Sexual		Total perpetration subscale	
	Mean ranks	p	Mean ranks	p	Mean ranks	p	Mean ranks	p	Mean ranks	p	Mean ranks	p
Physical												
No (N = 1077)	542.48	.001*	543.51	.021*	542.06	<.001*	543.13	<.001*	544.01	<.001*	541.85	<.001*
Yes (N = 14)	816.93		737.71		849.07		766.79		699.11		865.25	
Sexual												
No (N = 1089)	545.67	.418	545.24	.061	545.41	.138	545.51	<.001*	545.54	.019*	545.33	.100
Yes (N = 2)	724.00		959.50		867.25		810.50		795.25		912.00	

Note: The effect size for Mann–Whitney *U* test was small (PSest ≤ .56) in all statistically significant differences.

Abbreviations: *N*, sample size; *p*, *p* value (*$p < .05$).

TABLE A4 COSMIN risk of bias checklist. Evaluating the MSDV 2.0

Psychometric property	Score		
1. PROM development	Rater 1	Rater 2	Consensus
1.a PROM design	V	V	V
1.b Cognitive interview study or other pilot test	A	V	A
Total lowest score of items 1.a–1.b	A	V	A
2. Content validity	Rater 1	Rater 2	Consensus
2.a Asking patients about relevance	V	V	V
2.b Asking patients about comprehensiveness	V	V	V
2.c Asking patients about comprehensibility	V	V	V
2.d Asking professionals about relevance	V	V	V
2.e Asking professionals about comprehensiveness	V	V	V
Total Lowest score of items 2.a–2.e	V	V	V
3. Structural validity	Rater 1	Rater 2	Consensus
3.1 For CTT: Was exploratory or confirmatory factor analysis performed?	V	V	V
3.2 For IRT/Rasch: does the chosen model fit to the research question?	V	V	V
3.3 Was the sample size included in the analysis adequate?	V	V	V
3.4 Were there any other important flaws?	V	V	V
Total Lowest score of items 1–4	V	V	V
4. Internal consistency	Rater 1	Rater 2	Consensus
4.1 Was an internal consistency statistic calculated for each unidimensional (sub) scale separately?	V	V	V
4.2 For continuous scores: Was Cronbach's alpha or omega calculated?	V	V	V
4.3 For dichotomous scores: Was Cronbach's alpha or KR-20 calculated?	V	V	V
4.4 For IRT-based scores: Was standard error of the theta (SE [θ]) or reliability coefficient of estimated latent trait value (index of [subject or item] separation) calculated?	V	V	V
4.5 Were there any other important flaws?	V	V	V
Total Lowest score of items 1–5	V	V	V
5. Cross-cultural validity\measurement invariance	NT		
6. Reliability	NT		
7. Measurement error	NT		
8. Criterion validity	NT		
8.1 For continuous scores: Were correlations, or the area under the receiver operating curve calculated?	V	V	V
8.2 For dichotomous scores: Were sensitivity and specificity determined?	V	V	V
8.3 Were there any other important flaws?	V	V	V
Total lowest score of items 1–3	V	V	V
9. Hypotheses testing for construct validity			
9a. Comparison with other outcome measurement instruments (convergent validity)	Rater 1	Rater 2	Consensus
9.a.1 Is it clear what the comparator instrument(s) measure(s)?	V	V	V
9.a.2 Were the measurement properties of the comparator instrument(s) adequate?	A	V	A
9.a.3 Was the statistical method appropriate for the hypotheses to be tested?	V	V	V
9.a.4 Were there any other important flaws?	V	V	V
Total lowest score of items 1–4	A	V	A
9b. Comparison between subgroups (discriminative or known-groups validity)	NT		
10. Responsiveness	NT		

Score: V = very good; A = adequate; D = doubtful; I = inadequate; N = not applicable. Not rated (NT) = only those parts of the boxes need to be completed for which its psychometric property has been realized.